Optimality of Type I Orthogonal Arrays for Cross-over Models with Correlated Errors

by J. Kunert[1] and R.J. Martin[2]

[1] Fachbereich Statistik, Universität, D-44221 Dortmund, Germany.

[2] School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, U.K.

Abstract

We show that binary designs for cross-over models obtained from orthogonal arrays are universally optimal under correlated errors.

Some key words: Carry-over effect; Change-over design; Cross-over design; Dependent observations; Generalized least-squares; Optimal design; Orthogonal arrays; Repeated measurements design.

1. Introduction

We consider cross-over designs, often referred to as repeated measurements designs, with more than two treatments, where the within-block errors are assumed correlated with a known but arbitrary correlation structure. It is assumed for a design d that the measurement y_{dup} on unit (or subject) u at period p satisfies

$$y_{dup} = \tau_{d(u,p)} + \rho_{d(u,p-1)} + \alpha_p + \beta_u + e_{up},$$

where $\tau_{d(u,p)}$ is the direct effect of the treatment *d(u,p)*, assigned to the unit *u* at period *p* by the design *d*, $\rho_{d(u,p-1)}$ is the carry-over effect of the treatment received in the preceding period, α_p and β_u are the effects of the *p*-th period and the *u*-th unit respectively, and the $e_{up}, 1 \le u \le n, 1 \le p \le k$ are the errors, which have expectation 0. We assume that the errors between different units are uncorrelated, while the errors within each unit are correlated, all with the same known correlation structure, such that the covariance matrix of the errors within one unit equals $\sigma^2 S$, where S is a known ($k \times k$) matrix while σ^2 maybe unknown. We are looking for a design *d* which is optimal for the weighted least squares estimate, using *S*, of direct treatment contrasts, for given numbers of treatments *t*, periods *k* and units *n*.

It has been a long-standing theoretical problem to find optimal designs when there are carry-over effects. The additional treatment parameters make it impossible in general to consider the information matrix for the direct effects. Some restricted solutions to the problem are discussed in §2. In this note we are able to use a new result of Kushner (1997) to find, when $t \ge k$, binary designs which are optimal over a much wider class of designs.

2. Optimality of binary designs

In vector notation, with $Y = (y_{11}, ..., y_{1k}, ..., y_{nk})^T$, our model can be rewritten as

$$Y = T_d \tau + F_d \rho + (1_n \otimes I_k)\alpha + (I_n \otimes 1_k)\beta + e, \qquad (1)$$

where T_d is the ($nk \times t$) treatment design matrix, F_d is the ($nk \times t$) carry-over design matrix, $\otimes$ is the Kronecker product, I_n is the n x n identity matrix and 1_n is the n-vector of ones. The model assumes for the covariance-matrix of the errors that $Cov\ e = I_n \otimes S$.

As in Kunert (1991, formula (2)), the information matrix of a repeated measurements design d in model (1) can be written as

$$C_d = T_d^T \left(I_n \otimes S^{-\frac{1}{2}} \right) \omega^{\perp} \left\{ \left(I_n \otimes S^{-\frac{1}{2}} \right) \left[I_n \otimes 1_k, 1_n \otimes I_k, F_d \right] \right\} \left(I_n \otimes S^{-\frac{1}{2}} \right) T_d,$$

where for any a x b matrix M, we define $\omega^{\perp}(M) = I_a - M(M^T M)^- M^T$ (with $(M^T M)^-$ denoting a generalized inverse) and where $S^{-\frac{1}{2}}$ is a ($k \times k$) matrix with the property $S^{-\frac{1}{2}} S^{-\frac{1}{2}} = S^{-1}$.

To show a design is universally optimal, we use Kiefer's (1975) Proposition 1. It suffices to derive an upper bound for tr C_d and to show that there is a design d^* with a completely symmetric information matrix aB_t where $B_t = I_t - \frac{1}{t} 1_t 1_t^T$ for some scalar a, such that tr C_{d^*} equals this upper bound. Then d^* is universally optimum and every other design must have the same information matrix as *d** when it is optimal under any of the optimality criteria considered by Kiefer (1975).

As in Kunert (1991), it is easy to show that

$$C_d \le T_d^T\left(I_n \otimes S^{-\frac{1}{2}}\right)\omega^{\perp}\left(\left\{\left(I_n \otimes S^{-\frac{1}{2}}\right)\left[I_n \otimes 1_k, F_d B_t\right]\right\}\right)\left(I_n \otimes S^{-\frac{1}{2}}\right)T_d = \tilde{C}_d ,$$

where $A \le B$ means B-A is non-negative definite. Equality holds for any design which is uniform on the periods, i.e. where each treatment appears equally often in each period. Here $\tilde{C}_d$ is the information matrix in the model which leaves out α and replaces F_d by $F_d B_t$ in equation (1).

Let S* denote $S^{-1} - (1_k^T S^{-1} 1_k)^{-1} S^{-1} 1_k 1_k^T S^{-1} = S^{-1/2}\omega^{\perp}(S^{-1/2} 1_k) S^{-1/2}$. Then this upper bound of the information matrix can be written as

$$\tilde{C}_d = \tilde{C}_{d11} - \tilde{C}_{d12}\,\tilde{C}_{d22}^{-}\,\tilde{C}_{d12}^{T},$$

where $\tilde{C}_{d11} = T_d^T\left(I_n \otimes S^*\right)T_d$,

$$\tilde{C}_{d12} = T_d^T\left(I_n \otimes S^*\right)F_d B_t$$

and

$$\tilde{C}_{d22} = B_t F_d^T\left(I_n \otimes S^*\right)F_d B_t .$$

It is easy to come up with a candidate for an optimal design. It was shown by Martin and Eccleston (1997) that letting d^* be a type I orthogonal array $OA_I(n, k, t, 2)$, with the rows as periods and columns as units, we have

$$C_{d^*} = \tilde{C}_{d^*}$$

and that for all $i, j = 1,2$ we have $\tilde{C}_{d^*ij} = a_{ij} B_t$ for some appropriate number a_{ij}. So for the type I orthogonal array, C_{d^*} is completely symmetric and it only remains to show that d^* maximizes the trace of $\tilde{C}_d$.

However, the trace of $\tilde{C}_d$ for an arbitrary design d cannot be calculated, because it uses a generalized inverse of the matrix $\tilde{C}_{d22}$, which depends on the design d. Since there is no closed formula for tr $\tilde{C}_d$, the maximization of tr C_d in the case of a repeated measurements design appears to be difficult. There have been many attempts to circumvent this problem, namely:

(i) attention is restricted to a subset of designs for which $\tilde{C}_{d22}$ is fixed, see e.g. Cheng and Wu (1980),

(ii) in the case $k = 2$ there is a general solution for $\tilde{C}^{+}_{d22}$ (see Hedayat and Zhao, 1990),

(iii) in the case $t = 2$ $\tilde{C}_{d22}$ has rank 1 and is proportional to B_t, and there is a general formula for tr $\tilde{C}_d$, namely (see Kunert, 1991),

$$\text{tr}\,\tilde{C}_d = \text{tr}\,\tilde{C}_{d11} - (\text{tr}\,\tilde{C}_{d12})^2 / \text{tr}\,\tilde{C}_{d22}, \tag{2}$$

(iv) upper bounds for $\tilde{C}_{d22}$ are used (see e.g. Kunert, 1984).

None of these approaches are useful for the present situation: with correlated errors the set of all designs with the same $\tilde{C}_{d22}$ is very small, we have more than two treatments and more than two periods, and there is no upper bound for $\tilde{C}_{d22}$ which is really useful.

However, Kushner (1997, Lemma 5.1) introduced a new general bound for tr C_d which is given by (2), unless $\tilde{C}_{d22} = 0$, when it is tr $\tilde{C}_{d11}$.

In what follows we maximize the bound in (2), and show that d^* attains the maximum. To do this we restrict to the case $t \geq k$ and to binary designs, i.e. designs for which each unit receives each treatment at most once. Note that the type I orthogonal array d^* is a binary design.

Each of the $\tilde{C}_{dij}$ can be written as (see, e.g. Kunert, 1991),

$$\tilde{C}_{dij} = \sum_{u=1}^{n} \tilde{C}_{dij}^{(u)}$$

where e.g. $\tilde{C}_{d12}^{(u)} = T_{du}^T\, S * F_{du} B_t$ and the ($k \times t$) matrices T_{du}, F_{du} are the contributions of unit u to the treatment design matrix T_d, and the carry-over design matrix F_d, respectively. Now, for any binary design d each T_{du} can be written in the form $T_{du} = \left[I_k, 0_{(t-k)\times k}\right]\Pi_{du}$, where Π_{du} is a ($t \times t$) permutation matrix (see, e.g. Kunert and Utzig, 1993). Similarly, $F_{du} = \left[V, 0_{(t-k)\times k}\right]\Pi_{du}$, with the ($k \times k$) matrix

$$V = \begin{bmatrix} 0 & & & & & 0 \\ 1 & 0 & & & & \\ 0 & 1 & 0 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 0 & 1 & 0 & \\ 0 & & & 0 & 1 & 0 \end{bmatrix}.$$

It follows that for any binary design and for any unit u, we have

$$\begin{aligned}
\operatorname{tr} \tilde{C}_{d12}^{(u)} &= \operatorname{tr}\left(T_{du}^T\, S * F_{du} B_t\right) \\
&= \operatorname{tr}\left(\Pi_{du} B_t \Pi_{du}^T \begin{bmatrix} I_k \\ 0_{(t-k)\times k} \end{bmatrix} S * \left[V, 0_{k\times(t-k)}\right]\right) \\
&= \operatorname{tr}\left(B_t \begin{bmatrix} I_k \\ 0_{(t-k)\times k} \end{bmatrix} S * \left[V, 0_{k\times(t-k)}\right]\right) \\
&= c_{12}, \text{ say,}
\end{aligned}$$

which does not depend on the design d, but is the same for every binary design!

Similarly, $\operatorname{tr} \tilde{C}_{d11}^{(u)} = c_{11}$ and $\operatorname{tr} \tilde{C}_{d22}^{(u)} = c_{22}$ are the same for every binary design.

From the above results, it follows that for every binary design d we have

$$tr\, C_d \le nc_{11} - nc_{12}^2 / c_{22},$$

with equality holding for d*, since $C_{d^*} = \tilde{C}_{d^*}$. Hence we have shown that the type I orthogonal array d* is universally optimal for the estimation of direct effects over all binary designs with the same number of treatments, experimental units and periods. Note that this holds for any (known) correlation structure S.

3. Discussion

Although, under the stated conditions, we have obtained an optimal binary design for the model (1), Kushner (1997) has shown that using mixtures of sequences, some of which are non-binary, may be more efficient. However, the optimality of the binary design is important because binary designs are widely used and are acceptable to practitioners, and because the binary designs are optimal if the carry-over effects are absent and the errors are uncorrelated.

The above result can be generalized easily to other cross-over models and situations. Some examples are: if single within-unit differencing is necessary for stationarity, and in the circular case when there is a preperiod.

Acknowledgement

The authors wish to thank H.B. Kushner for prior notice of his work.

References:

Cheng, C. S. and Wu, C. F. (1980): Balanced repeated measurements designs. *Annals of Statistics* 8, 1272 - 1283. Corrigendum 11 (1983), 349.

Hedayat, A. and Zhao, W. (1990): Optimal two-period repeated measurements designs. *Annals of Statistics* 18, 1805 - 1816.

Kiefer, J. (1975): Construction and optimality of generalized Youden designs. In *'A survey of statistical design and linear models'*, Ed. J. N. Srivastava, pp. 333-53. North Holland: Amsterdam.

Kunert, J. (1984): Optimality of balanced uniform repeated measurements designs. *Annals of Statistics* 12, 1006 - 1017.

Kunert, J. (1991): Cross-over designs for two treatments and correlated errors. *Biometrika* 78, 315 - 324

Kunert, J. and Utzig, B. (1993): Estimation of variance in cross-over designs. *Journ. Roy. Statist. Soc. B* 55, 919 - 927.

Kushner, H. B. (1997): Optimal repeated measurements designs: the linear optimality equations. *Annals of Statistics* 25, 000-000.

Martin, R. J. and Eccleston, J. A. (1997): Variance-balanced change-over designs for dependent observations. *University of Sheffield Research Report*.